

NASA/CR-2000-210625
ICASE Report No. 2000-47



Enhanced Accuracy by Post-processing for Finite Element Methods for Hyperbolic Equations

Bernardo Cockburn and Mitchell Luskin
University of Minnesota, Minneapolis, Minnesota

Chi-Wang Shu
Brown University, Providence, Rhode Island

Endre Süli
Oxford University Computing Laboratory, Oxford, United Kingdom

ICASE
NASA Langley Research Center
Hampton, Virginia
Operated by Universities Space Research Association



National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NAS1-97046

December 2000

Form SF298 Citation Data

Report Date <i>("DD MON YYYY")</i> 00DEC2000	Report Type N/A	Dates Covered (from... to) <i>("DD MON YYYY")</i>
Title and Subtitle Enhanced Accuracy by Post-processing for Finite Element Methods for Hyperbolic Equations	Contract or Grant Number	
	Program Element Number	
Authors Chi-Wang Shu, Endre Süli	Project Number	
	Task Number	
	Work Unit Number	
Performing Organization Name(s) and Address(es) ICASE NASA Langley Research Center Hampton, Virginia	Performing Organization Number(s) NASA/CR-2000-210625 ICASE Report No. 2000-47	
Sponsoring/Monitoring Agency Name(s) and Address(es)	Monitoring Agency Acronym	
	Monitoring Agency Report Number(s)	
Distribution/Availability Statement Approved for public release, distribution unlimited		
Supplementary Notes		
Abstract Abstract. We consider the enhancement of accuracy, by means of a simple post-processing technique, for finite element approximations to transient hyperbolic equations. The post-processing is a convolution with a kernel whose support has measure of order one in the case of arbitrary unstructured meshes; if the mesh is locally translation invariant, the support of the kernel is a cube whose edges are of size of the order of the mesh size only. For example, when polynomials of degree k are used in the discontinuous Galerkin (DG) method, and the exact solution is globally smooth, the DG method is of order $k + 1 = 2$ in the L_2 norm, whereas the post-processed approximation is of order $2k + 1$; if the exact solution is in L_2 only, in which case no order of convergence is available for the DG method, the post-processed approximation converges with order $k + 1 = 2$ in L_2 over a subdomain on which the exact solution is smooth. Numerical results displaying the sharpness of the estimates are presented.		
Subject Terms		
Document Classification unclassified	Classification of SF298 unclassified	

Classification of Abstract unclassified	Limitation of Abstract unlimited
Number of Pages 30	

ENHANCED ACCURACY BY POST-PROCESSING FOR FINITE ELEMENT METHODS FOR HYPERBOLIC EQUATIONS

BERNARDO COCKBURN*, MITCHELL LUSKIN†, CHI-WANG SHU‡, AND ENDRE SÜLI§

Abstract. We consider the enhancement of accuracy, by means of a simple post-processing technique, for finite element approximations to transient hyperbolic equations. The post-processing is a convolution with a kernel whose support has measure of order one in the case of arbitrary unstructured meshes; if the mesh is locally translation invariant, the support of the kernel is a cube whose edges are of size of the order of the mesh size only. For example, when polynomials of degree k are used in the discontinuous Galerkin (DG) method, and the exact solution is globally smooth, the DG method is of order $k + 1/2$ in the L^2 norm, whereas the post-processed approximation is of order $2k + 1$; if the exact solution is in L^2 only, in which case no order of convergence is available for the DG method, the post-processed approximation converges with order $k + 1/2$ in L^2 over a subdomain on which the exact solution is smooth. Numerical results displaying the sharpness of the estimates are presented.

Key words. post-processing, finite element methods, hyperbolic problems

Subject classification. Applied and Numerical Mathematics

1. Introduction. In this paper, we consider general finite element methods for time-dependent *linear* hyperbolic systems of the form

$$\begin{aligned} u_t + \sum_{j=1}^d A_j u_{x_j} + A_0 u &= 0, & (x, t) \in \mathbb{R}^d \times (0, T], \\ u(x, 0) &= u_0(x), & x \in \mathbb{R}^d, \end{aligned}$$

where $\{A_j\}_{j=1}^d$ are real, constant coefficient $m \times m$ matrices such that $\sum_{j=1}^d A_j \xi_j$ has real eigenvalues and a complete set of linearly independent eigenvectors for all $\xi \in \mathbb{R}^d$, and the function u has range in \mathbb{R}^m . Our aim in this paper is to show how to exploit the *inherently oscillatory nature* of numerical solutions to this problem computed by means of finite element methods to enhance the quality of the approximation. This enhancement is achieved by post-processing the approximate solution *only once*, at the very end of the computation, at $t = T$. The post-processing considered here is completely independent of the partial differential equation under consideration and can be performed for entirely arbitrary triangulations; however, it takes a particularly simple and computationally efficient form when the triangulation is *locally translation invariant*.

*School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (cockburn@math.umn.edu). Supported in part by NSF Grant DMS-9807491 and by the University of Minnesota Supercomputing Institute.

†School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (luskin@math.umn.edu). Supported in part by NSF Grant DMS 95-05077, by AFOSR Grant F49620-98-1-0433, by ARO Grant DAAG55-98-1-0335, by the Institute for Mathematics and its Applications, and by the Minnesota Supercomputing Institute.

‡Division of Applied Mathematics, Brown University, Providence, RI 02912 (shu@cfdm.brown.edu). Supported in part by ARO Grants DAAG55-97-1-0318 and DAAD19-00-1-0405, NSF Grant DMS-9804985, NASA Langley Grant NAG-1-2070 and Contract NAS1-97046 while this author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23681-2199, and AFOSR Grant F49620-99-1-0077.

§Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, U.K. (Endre.Suli@comlab.ox.ac.uk). This author is grateful to the Institute for Mathematics and its Applications at the University of Minnesota and the University of Minnesota Supercomputing Institute for their generous support.

To illustrate the basic idea, let us consider the following simple model problem:

$$u_t + u_x = 0, \quad \text{in } (0, 1) \times (0, T), \quad u(x, 0) = \sin(2\pi x) \quad \text{for } x \in (0, 1),$$

subject to periodic boundary conditions, and let us compute an approximation U to its solution u by using the discontinuous Galerkin (DG) method with piecewise polynomials of degree one over *uniform* grids of spacing h . We also consider the post-processed approximation $U^* = K_h^{4,2} \star U$, where the convolution kernel $K_h^{4,2}(x) = \frac{1}{h} K^{4,2}(x/h)$ is defined by

$$K^{4,2}(y) = -\frac{1}{12}\psi^{(2)}(y-1) + \frac{7}{6}\psi^{(2)}(y) - \frac{1}{12}\psi^{(2)}(y+1),$$

where $\psi^{(2)}$ is the B-spline obtained by convolving the characteristic function $\psi^{(1)} = \chi$ of the interval $(-1/2, 1/2)$ with itself once. In Fig. 1.1 we display, for $T = 0.1$ and $h = 1/10$ and $h = 1/20$, the errors $x \mapsto u(T, x) - U(T, x)$ and $x \mapsto u(T, x) - U^*(T, x)$. The time-step was chosen so small that the overall accuracy of the method is dominated by the spatial error. We note the oscillatory nature of the error $x \mapsto u(T, x) - U(T, x)$ typical of finite element methods and the apparent superconvergence of the numerical solution at the two Gauss-Radau points, a fact discovered in 1995 by Adjerid, Aifia, and Flaherty [2]; see also their recent work [1]. In contrast with this behavior, we observe the complete absence of oscillations from the error $u(T) - U^*(T)$. This shows that convolving the approximate solution U with the kernel $K_h^{4,2}$ *filters out* the numerical oscillations around the exact solution. Moreover, the result of such a filtering is a new approximation U^* that converges faster to u than U . Indeed, in Fig. 1.2, we display the functions $x \mapsto \log(|u(T, x) - U(T, x)|)$, for $h = 1/10, 1/20, 1/40$ and $1/80$; we observe that each time h is halved, the maximum of $x \mapsto |u(T, x) - U^*(T, x)|$ is divided by a factor not less than eight. This indicates that the post-processed approximation is at least third-order convergent; the original approximate solution U exhibits only second-order convergence.

In Figs. 1.3 and 1.4 we repeat the above experiment using polynomials of degree two. Again we observe the oscillatory nature of the approximation and the superconvergence at the three Gauss-Radau points in Fig. 1.3 (top), and that the oscillations are filtered out upon convolution in Fig. 1.3 (bottom). This time, the convolution kernel $K_h^{6,3}(x) = \frac{1}{h} K^{6,3}(x/h)$ is defined by

$$\begin{aligned} K^{6,3}(y) = & \frac{37}{1920}\psi^{(4)}(y-2) - \frac{97}{480}\psi^{(4)}(y-1) - \frac{437}{320}\psi^{(4)}(y) \\ & - \frac{97}{480}\psi^{(4)}(y+1) + \frac{37}{1920}\psi^{(4)}(y+2), \end{aligned}$$

where $\psi^{(4)}$ is the B-spline obtained by convolving the characteristic function $\psi^{(1)} = \chi$ of the interval $(-1/2, 1/2)$ with itself three times. In Fig. 1.4, we see that each time h is halved, the maximum error decreases by a factor not less than thirty two. This shows that the error in the post-processed approximation is of fifth order.

In connection with this fact, we note here that in 1996 Lowrie [16] found analytical and numerical evidence that when polynomials of degree k are used, a ‘component of the error’ of the DG method converges with order $2k + 1$ in the L^2 norm; this fact stands in striking contrast with convergence of order $k + 1/2$ for the underlying DG approximation ($k + 1$ for the one-dimensional case and special grids in several space dimensions). In this paper, we provide a firm mathematical basis for this observation, and show how to compute the superconvergent approximation U^* by a simple post-processing technique which is independent of the equation and of the numerical method.

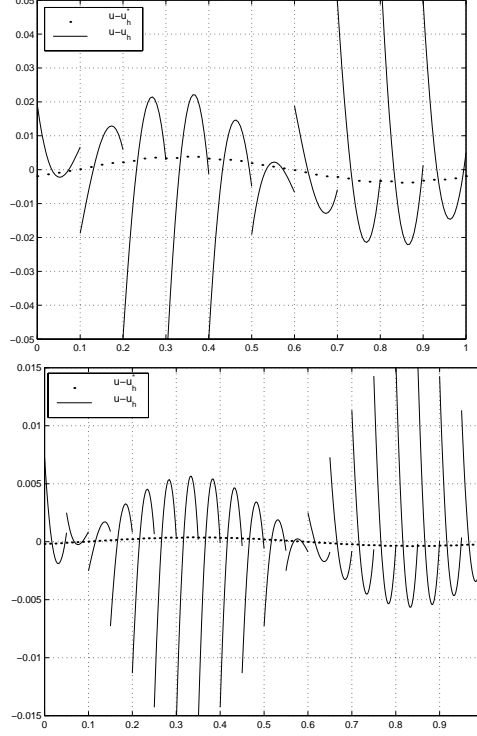


FIG. 1.1. The errors $u - U$ (solid line) and $u - U^*$ (dots) at $T = 0.1$ for $h = 1/10$ (top) and $h = 1/20$ (bottom). The function u is the smooth exact solution, U is the approximation given by the DG method with polynomials of degree one, and $U^* = K_h^{4,2} \star U$.

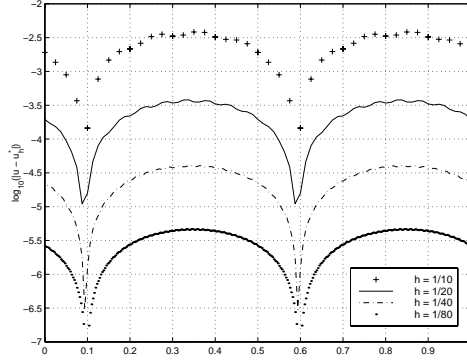


FIG. 1.2. The errors $\log(|u - U^*|)$ at $T = 0.1$ for $h = 1/10$ (top), $h = 1/20$, $h = 1/40$, and $h = 1/80$ (bottom). Each time h is halved, the maximum error decreases by a factor not less than 8; the order of convergence is, therefore, not less than 3.

The paper is organized as follows. In Section 2, we present a brief account of the development of the ideas behind this paper. In Section 3, we state and discuss our main theoretical results, and in Section 4 we present their proofs. In Section 5, we display numerical experiments which not only verify our theoretical results but also indicate how this kind of post-processing can be applied to convection-diffusion and non-linear problems. We conclude, in Section 6, with some remarks.

2. A brief overview of the development of post-processing techniques. In order to introduce the basic ideas of our work and to put them into proper perspective, we briefly review the development of

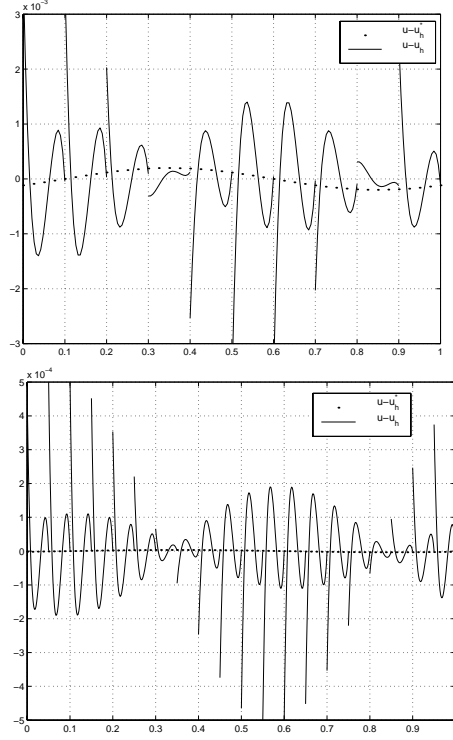


FIG. 1.3. The errors $u - U$ (solid line) and $u - U^*$ (dots) at $T = 0.1$ for $h = 1/10$ (top) and $h = 1/20$ (bottom). The function u is the smooth exact solution, U is the approximation given by the DG method with polynomials of degree two, and $U^* = K_h^{6,3} \star U$.

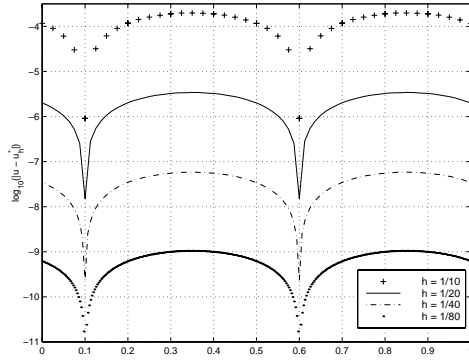


FIG. 1.4. The errors $\log(|u - U^*|)$ at $T = 0.1$ for $h = 1/10$ (top), $h = 1/20$, $h = 1/40$, and $h = 1/80$ (bottom). Each time h is halved, the maximum error decreases by a factor not less than 32; the order of convergence is, therefore, not less than 5.

post-processing techniques devised to improve the quality of numerical approximations. For further details, the reader should consult the monograph of Wahlbin [21] on superconvergence in Galerkin finite element methods.

2.1. Finite difference and spectral methods for hyperbolic problems. In 1977, Majda and Osher [18] considered formally high-order accurate dissipative difference schemes for hyperbolic problems. They studied a one-dimensional model problem of a two-by-two hyperbolic system whose characteristics are parallel to $x = \pm t$; the initial condition is a step function whose discontinuity is located at the origin. Majda

and Osher showed that the rate of convergence on the region between the characteristics issuing from the origin, $|x/t| < 1 - \delta^2$, is independent of the numerical scheme. They pointed out that in 1962 Fedorenko [11] and in 1969 Apelkrans [3] displayed numerical evidence that the order of convergence had to be one. However, by selecting a suitable approximation of the initial datum, Majda and Osher showed that the order of convergence can be increased to two. Moreover, they found that they could recover the full formal order of accuracy of the scheme on the region $|x/t| < 1 - \delta^2$ provided they preprocessed the initial data in an appropriate way. In 1986, Johnson and Pitkäranta [14] used a similar idea in the analysis of the DG method for linear hyperbolic problems. The question of post-processing the initial data is considered in the book of Brenner, Thomée and Wahlbin [5]; see also the work of Jovanović, Ivanović and Süli [15] concerning the use of convolution mollifiers with B-spline kernels for second-order hyperbolic boundary value problems with non-smooth data.

In 1978, Mock and Lax [19] showed that for a difference scheme of any formal order of accuracy μ , for linear hyperbolic systems, the *moments* of the exact solution converge with order μ provided that, again, the initial data was suitably preprocessed. This result holds even if the exact solution contains discontinuities. They also showed how to post-process the approximate solution by a simple convolution to enhance its accuracy over regions of smoothness of the exact solution: if the solution was sufficiently smooth locally, they could obtain nearly the full order of convergence μ provided that the support of the kernel was of order almost one. This seems to have been the first instance when the ideas of (i) preprocessing the initial data, (ii) obtaining error estimates for the moments, and (iii) post-processing the approximation, appear clearly delineated.

Later, in 1985, Gottlieb and Tadmor [12], motivated by the work of Mock and Lax [19], found a spectrally accurate post-processing kernel for spectral methods; see also the 1978 paper by Majda, McDonough and Osher [17]. Again, the full spectral accuracy could be recovered by using a convolution; the measure of the support of the kernel had to be of order one.

2.2. Finite element methods for elliptic problems. Quite independently of the developments reviewed above, in 1977 Bramble and Schatz [4] considered linear elliptic problems and showed how to post-process the finite element solution by means of a simple convolution to enhance the quality of the approximation. They showed that the order of convergence could be doubled if the exact solution was locally smooth. It is important to point out that, just like Mock and Lax, Bramble and Schatz proved a negative-order norm error estimate (an error estimate of the moments in Mock and Lax's terminology) and then showed how to use it to enhance the approximation by a convolution. However, unlike Mock and Lax's convolution kernel, for locally translation invariant grids the Bramble–Schatz kernel has support in a cube whose diameter is of order h only; this fact represents a considerable advantage from the computational point of view.

Also in 1977, Thomée [20] extended the work of Bramble and Schatz [4] to include superconvergence of the derivatives and gave an elegant proof of their approximation results by using Fourier analysis.

An application of the Bramble and Schatz technique to the simulation of miscible displacement was devised and analyzed by Douglas [9]; other applications can be found in the book of Wahlbin [21].

2.3. The main ideas. In this paper, we apply the ideas of Mock and Lax [19] and Bramble and Schatz [4] to enhance the accuracy of finite element approximations to hyperbolic problems by post-processing. We proceed as follows. First, we obtain an estimate of the error between the analytical solution u and the post-processed numerical approximation U in terms of negative-order Sobolev norms of $u - U$. This result does not depend on the partial differential equation under consideration or on the numerical scheme. Next,

we obtain negative-order norm a priori estimates for the error between the exact solution of a hyperbolic problem and its finite element approximation U . The final error estimate is then obtained by combining the above bounds.

3. The results. In this section, we present and discuss our main theoretical results.

3.1. An approximation result. We begin by presenting a result that relates negative-order norm a priori estimates of the difference between u and an arbitrary approximation U for u to L^2 -error estimates of the difference between u and the post-processed counterpart U .

Let us recall the definition of a negative-order Sobolev norm on an open set $\Omega \subset \mathbb{R}^d$. We denote by $\|u\|_{0,\Omega}$ the standard L^2 -norm of u on Ω . For any natural number ℓ , we consider the norm and seminorm of the Sobolev space $H^\ell(\Omega)$, defined by

$$\|u\|_{\ell,\Omega} = \left\{ \sum_{|\alpha| \leq \ell} \|D^\alpha u\|_{0,\Omega}^2 \right\}^{1/2}, \quad |u|_{\ell,\Omega} = \left\{ \sum_{|\alpha| = \ell} \|D^\alpha u\|_{0,\Omega}^2 \right\}^{1/2}.$$

Sobolev norms and seminorms for vector-valued functions from $H^\ell(\Omega, \mathbb{R}^m)$ are defined analogously and are denoted by the same symbol as in the scalar case. We then define the negative order Sobolev norm $\|\cdot\|_{-\ell,\Omega}$, $\ell \geq 1$, by

$$\|u\|_{-\ell,\Omega} = \sup_{\phi \in \mathcal{C}_0^\infty(\Omega)} \frac{\int_\Omega u(x) \phi(x) dx}{\|\phi\|_{\ell,\Omega}}.$$

Negative-order norms can be used to detect the oscillations of a function around zero. For example, for $\Omega = (-1, 1)$, $\ell \geq 1$ and $u_N(x) = \sin(2\pi N x)$, a simple computation gives $\|u_N\|_{-\ell,\Omega} = 1/(2\pi N)^\ell$, indicating that u_N oscillates about zero in a very regular manner.

Next, we describe the type of post-processing to be considered following Bramble and Schatz [4]. We post-process the approximate solution by convolving it with a kernel $K_H^{\nu,\ell}(x) = K^{\nu,\ell}(x/H)/H^d$ which has to satisfy three properties; the first of these is that $K^{\nu,\ell}$ has *compact support*. The second is that it reproduces polynomials p of degree $\nu - 1$ by convolution, that is,

$$K^{\nu,\ell} \star p = p.$$

This is the type of kernel used by Mock and Lax [19]. The kernels used by Bramble and Schatz [4] which we shall next describe have the further property that they are linear combinations of B-splines. Let χ be the characteristic function of the interval $(-1/2, 1/2)$ and let δ denote the Dirac distribution concentrated at $x = 0$. Then, we define recursively the functions $\psi^{(i)}$ as follows:

$$\psi^{(0)} = \delta, \quad \psi^{(n+1)} = \psi^{(n)} \star \chi, \text{ for } n \geq 0,$$

and, given an arbitrary multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$ and $y = (y_1, \dots, y_d) \in \mathbb{R}^d$, we set

$$\psi^{(\alpha)}(y) = \psi^{(\alpha_1)}(y_1) \dots \psi^{(\alpha_d)}(y_d).$$

We also set $\mathbf{1} = (1, \dots, 1)$. The third, and final, property of the kernels considered here is that they are of the form

$$K^{\nu,\ell\mathbf{1}}(y) = \sum_{\gamma \in \mathbb{Z}^d} k_\gamma^{\nu,\ell\mathbf{1}} \psi^{(\ell\mathbf{1})}(y - \gamma), \quad (3.1)$$

where $k_\gamma^{\nu, \ell 1} \in \mathbb{R}$. Note that since the support of $K^{\nu, \ell 1}$ has been assumed compact, there are only finitely many non-zero coefficients $k_\gamma^{\nu, \ell 1}$ in this sum.

The imposition of these hypotheses is motivated by the following observations: the compactness of the support of the convolution kernel is advantageous from the computational point of view; the second property ensures that accuracy of order ν is not destroyed by post-processing; the third property allows us to express derivatives of the convolution with the kernel in terms of simple difference quotients. Indeed, it is very easy to verify that for multi-indices α and β such that $\beta_i \geq \alpha_i$ for $i = 1, \dots, d$, we have

$$D^\alpha (\psi_H^{(\beta)} \star v) = \psi_H^{(\beta-\alpha)} \star \partial_H^\alpha v, \quad (3.2)$$

where $\psi_H^{(\alpha)}(x) = \psi^{(\alpha/H)} / H^d$,

$$\partial_H^\alpha := \partial_{H,1}^{\alpha_1} \dots \partial_{H,d}^{\alpha_d} \quad \text{and} \quad \partial_{H,j} v(x) = \frac{1}{H} (v(x + \frac{1}{2} H e_j) - v(x - \frac{1}{2} H e_j)).$$

This fact can then be exploited in the finite element framework, as will be seen later. We are now ready to state an approximation result which shows that local smoothness of u on the one hand and negative-order norm estimates of divided differences of the error $u - U$ on the other lead to a local bound on $u - K_H^{\nu, \ell 1} \star U$ in the L^2 -norm.

THEOREM 3.1 (Bramble and Schatz[4]). *Let ν and ℓ be two natural numbers. Suppose, further, that $K_H^{\nu, \ell 1}(x) = K^{\nu, \ell 1}(x/H)/H^d$ where $K^{\nu, \ell 1}$ is a function of compact support which reproduces polynomials of degree $\nu - 1$ by convolution, and which is the linear combination of B-splines, as in (3.1). Let U be a function in $L^2(\Omega_1)$, where Ω_1 is an open set in \mathbb{R}^d , and let u be a function in $H^\nu(\Omega_1)$. Let Ω_0 be an open set in \mathbb{R}^d such that $\Omega_0 + 2 \operatorname{supp}(K_H^{\nu, \ell 1}) \subset \subset \Omega_1$ for all $H \leq H_0$. Then, for $H \leq H_0$, we have*

$$\|u - K_H^{\nu, \ell 1} \star U\|_{0, \Omega_0} \leq \frac{H^\nu}{\nu!} C_1 |u|_{\nu, \Omega_1} + C_1 C_2 \sum_{|\alpha| \leq \ell} \|\partial_H^\alpha (u - U)\|_{-\ell, \Omega_1},$$

where $C_1 = \sum_{\gamma \in \mathbb{Z}^d} |k_\gamma^{\nu, \ell 1}|$ and C_2 depends solely on Ω_0 , Ω_1 , d , ν , and ℓ .

To illustrate the importance of this result, let us assume that there exist real numbers $\mu \geq 0$ and $a \in [0, \ell]$ such that, for all $H \leq H_0$,

$$\sum_{|\alpha| \leq \ell} \|\partial_H^\alpha (u - U)\|_{-\ell, \Omega_1} \leq C_3 h^\mu H^{-a}. \quad (3.3)$$

Note that the number a measures how well it is possible to estimate the negative-order norm of the divided differences of $u - U$. In the worst case, $a = \ell$; this is the case treated by Mock and Lax [19]. In the finite element framework, however, it is possible to take a to be different from ℓ , as Bramble and Schatz [4] showed for second-order elliptic problems.

Inserting the inequality (3.3) in the inequality of Theorem 3.1, we get

$$\begin{aligned} \|u - K_H^{\nu, \ell 1} \star U\|_{0, \Omega_0} &\leq \frac{H^\nu}{\nu!} C_1 |u|_{\nu, \Omega_1} + C_1 C_2 C_3 h^\mu H^{-a} \\ &\leq C_1 \max\{|u|_{\nu, \Omega_1}/\nu!, C_2 C_3\} (H^\nu + h^\mu H^{-a}). \end{aligned}$$

If we now define \hat{H} to be the solution of the equation $H^\nu = h^\mu H^{-a}$, we obtain the following result.

COROLLARY 3.2. *Let the hypotheses of Theorem 3.1 hold, and suppose that (3.3) is valid. Then, for $\hat{H} = h^{\mu/(\nu+a)} \leq H_0$, we have*

$$\|u - K_{\hat{H}}^{\nu, \ell 1} \star U\|_{0, \Omega_0} \leq C h^{\theta \mu},$$

where $C = 2 C_1 \max\{|u|_{\nu, \Omega_1} / \nu!, C_2 C_3\}$ and $\theta = \nu / (\nu + a)$.

Note that in the worst possible case, that is when $a = \ell$, this implies that

$$\|u - K_{\hat{H}}^{\nu, \ell 1} \star U\|_{0, \Omega_0} \leq C h^{\theta \mu},$$

with $\theta = \nu / (\nu + \ell) < 1$. The only possibility we then have for raising the order of convergence is to hope that the function u is very smooth so that we can choose ν large and positive. Unfortunately, even if this were actually possible, the support of the convolution kernel would be contained in a cube whose diameter is of order $\hat{H} = h^{\mu/(\nu+\ell)}$ which converges to a quantity of order one as ν increases to infinity; this in turn renders the evaluation of the convolution computationally inefficient.

On the other hand, in the best possible case (that is when $a = 0$), taking $\nu = \mu$ would permit choosing $\theta = 1$, $\hat{H} = h$ and we would then have

$$\|u - K_{\hat{H}}^{\nu, \ell 1} \star U\|_{0, \Omega_0} \leq C h^{\mu}.$$

In other words, for $a = 0$ we obtain the *same* order of convergence for $u - K_{\hat{H}}^{\nu, \ell 1} \star U$ in the local L^2 norm as that of the local negative-order norm error estimate in (3.3). Moreover, this is achieved by using a convolution kernel whose support is contained in a cube whose diameter is of the order $\hat{H} = h$ only; this renders the evaluation of the convolution a very fast computation. The examples shown in the Introduction correspond to this case with $\nu = \mu = 2k + 1$, where k is the degree of polynomials in the discontinuous Galerkin method.

3.2. Negative-order norm error estimates for finite element methods.

3.2.1. The weak solution. As stated in the Introduction, we consider the following Cauchy problem:

$$u_t + \sum_{j=1}^d A_j u_{x_j} + A_0 u = 0, \quad (x, t) \in \mathbb{R}^d \times (0, T], \quad (3.4)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}^d, \quad (3.5)$$

To make the presentation of the ideas as simple as possible, we reduce unessential technicalities by assuming that the matrices in the equation (3.4) are independent of time and space and by taking the initial data to be 1-periodic in each of the coordinate directions x_i , $i = 1, \dots, d$, and we seek a solution to the above problem which is 1-periodic in each coordinate direction.

We suppose that the system of equations (3.4) is strongly hyperbolic, that is, there exists a family of real $m \times m$ matrices $\{S(\xi) : \xi \in \mathbb{R}^d\}$ and a constant $K > 0$ such that

$$S(\xi) \left(\sum_{j=1}^d A_j \xi_j \right) S^{-1}(\xi)$$

is a diagonal matrix for all $\xi \in \mathbb{R}^d$, and

$$\sup_{|\xi|=1} (\|S(\xi)\| + \|S^{-1}(\xi)\|) \leq K. \quad (3.6)$$

Letting $I = (0, 1)^d$, the weak solution, $u(x, t)$, of (3.4) satisfies

$$(u, \varphi)_I(t) = (u_0, \varphi(0))_I + \int_0^t (u, \varphi_t + \sum_{j=1}^d A_j^* \varphi_{x_j} - A_0^* \varphi)_I d\tau \quad (3.7)$$

for all $\varphi \in C^\infty([0, T]; H_{\text{per}}^1(\mathbb{R}^d, \mathbb{R}^m))$ and $t \in [0, T]$ where A_j^* is the transpose of A_j and in the above equation and below

$$(u, \varphi)_I(t) = \int_I u(x, t) \varphi(x, t) dx.$$

Here, $H_{\text{per}}^1(\mathbb{R}^d, \mathbb{R}^m)$ denotes the Sobolev space of 1-periodic functions defined as follows. Let $\mathcal{C}_{\text{per}}^\infty(\mathbb{R}^d, \mathbb{R}^m)$ be the subset of $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^m)$ of 1-periodic functions. We then define $H_{\text{per}}^1(\mathbb{R}^d, \mathbb{R}^m)$ as the closure of $\mathcal{C}_{\text{per}}^\infty(\mathbb{R}^d, \mathbb{R}^m)$ for the $H^1(I, \mathbb{R}^m)$ -norm.

It follows from (3.6) that the problem (3.7) is well posed in

$$L_{\text{per}}^2(\mathbb{R}^d; \mathbb{R}^m) = \{f \in L_{\text{loc}}^2(\mathbb{R}^d; \mathbb{R}^m) : f(x + \alpha) = f(x) \text{ for all } x \in \mathbb{R}^d, \alpha \in \mathbb{Z}^d\}$$

with respect to the norm $\|\cdot\|_{L^2(I)}$; see Theorem 6.3.2. on p. 219 of [13].

3.2.2. The finite element methods. Next, we describe the class of finite element approximations to (3.4). It includes the standard Galerkin method, the Galerkin method with artificial diffusion and the discontinuous Galerkin method. With slight modifications we could have easily included, for example, the streamline diffusion method and the stabilized discontinuous Galerkin method; however, in order to avoid unnecessary technical complications, we have chosen not to consider these.

Let $\mathcal{T}_h = \{K\}$ be a regular triangulation of \mathbb{R}^d , invariant under translations by $\alpha \in \mathbb{Z}^d$, whose elements K are open and have diameter h_K less than or equal to h . It will be assumed throughout that each $K \in \mathcal{T}_h$ is contained either in I or in $\mathbb{R}^d \setminus I$. For a nonnegative integer k , we associate with the triangulation \mathcal{T}_h the broken Sobolev space

$$H_{\text{per},h}^k(\mathbb{R}^d; \mathbb{R}^m) = \Pi_{K \in \mathcal{T}_h} H^k(K; \mathbb{R}^m) \cap L_{\text{per}}^2(\mathbb{R}^d; \mathbb{R}^m).$$

For $k = 0$, we shall write $L_{\text{per},h}^2(\mathbb{R}^d; \mathbb{R}^m) = H_{\text{per},h}^0(\mathbb{R}^d; \mathbb{R}^m)$. We then consider two finite element subspaces \mathcal{M}_h and \mathcal{N}_h of $H_{\text{per},h}^1(\mathbb{R}^d; \mathbb{R}^m)$, and the broken L_2 inner product $(\cdot, \cdot)_h$ defined on $L_{\text{per},h}^2(\mathbb{R}^d; \mathbb{R}^m) \times L_{\text{per},h}^2(\mathbb{R}^d; \mathbb{R}^m)$ by

$$(W, \chi)_h = \sum_{K \in \mathcal{T}_{h,I}} (W, \chi)_K, \quad (3.8)$$

where $\mathcal{T}_{h,I} = \{K \in \mathcal{T} : K \subset I\}$.

We define the finite element approximation $U : [0, T] \rightarrow \mathcal{M}_h$ as the solution to

$$(U_t(t), \chi)_h + B(U(t), \chi) = 0, \quad \chi \in \mathcal{N}_h, \quad (3.9)$$

$$U(0) = P_h u_0, \quad (3.10)$$

where $B(\cdot, \cdot)$ is a bilinear form defined on $\mathcal{M}_h \times H_{\text{per},h}^1(\mathbb{R}^d; \mathbb{R}^m)$, and the operator $P_h : L_{\text{per}}^2(\mathbb{R}^d; \mathbb{R}^m) \rightarrow \mathcal{M}_h$ is the orthogonal projection in the norm of $L^2(I)$.

In Table 3.1, we describe different choices of the form B that give rise to different finite element methods; in each of these $\mathcal{M}_h = \mathcal{N}_h$, although this need not be the case in general.

The operator \mathcal{A} and the bilinear form $\langle \cdot, \cdot \rangle_h$ that appear in Table 3.1 are defined as follows:

$$\begin{aligned} \mathcal{A}\chi &= \sum_{j=1}^d A_j \chi_{x_j} + A_0 \chi, \\ \mathcal{A}^* \chi &= - \sum_{j=1}^d A_j^* \chi_{x_j} + A_0^* \chi, \\ \langle \mathcal{A}U, \chi \rangle_h &= \sum_{K \in \mathcal{T}_{h,I}} \sum_{e \in \partial K} \langle \mathcal{A} \cdot n \hat{U}, \eta \rangle_e, \end{aligned}$$

TABLE 3.1
Examples of finite element methods.

Method	$\mathcal{M}_h \subset \mathcal{C}^0$	$B(U, \eta)$
Standard Galerkin (SG)	yes	$(\mathcal{A}U, \eta)_I$
SG with artificial diffusion	yes	$(\mathcal{A}U, \eta)_I + h^\gamma (\nabla U, \nabla \eta)_I, \gamma \geq 1,$
Discontinuous Galerkin	no	$(U, \mathcal{A}^* \eta)_h + \langle \mathcal{A}U, \eta \rangle_h$

where $\mathcal{A} \cdot n = A_1 n_1 + \dots + A_d n_d$, $n = (n_1, \dots, n_d)$ is the unit outward normal vector to K on $e \subset \partial K$, and \widehat{U} is the numerical flux of the DG method defined as follows. Given an element K and a face $e \in \partial K$, let us denote by $K_e \in \mathcal{T}_h$ the element sharing the edge e with K and denote by U_K and U_{K_e} the traces of U on e from K and K_e , respectively. We compute the $m \times m$ diagonal matrix $\text{diag}(\lambda_1, \dots, \lambda_m) = S(n) (\mathcal{A} \cdot n) S^{-1}(n)$ and set $V = S^{-1}(n) U$ and

$$\widehat{V}_j = \begin{cases} (V_K)_j & \text{if } \lambda_j > 0, \\ (V_{K_e})_j & \text{otherwise.} \end{cases}$$

The numerical flux is defined as follows:

$$\widehat{U} = S(n) \widehat{V}. \quad (3.11)$$

3.2.3. The negative-order error estimate. We now give sufficient conditions for the finite element method which ensure that, for a given time T , our approximate solution, $U(T)$, converges with high order in a negative-order norm over a given subdomain $\Omega_0 \subset\subset I$ to the weak solution $u(T)$. Given that $l \geq 0$, we wish to estimate

$$\|u(T) - U(T)\|_{-\ell, \Omega_0} = \sup_{\Phi \in \mathcal{C}_0^\infty(\Omega_0)} \frac{(u(T) - U(T), \Phi)}{\|\Phi\|_{\ell, \Omega_0}}.$$

We begin by considering the solution to the dual problem: Find a function φ such that $\varphi(\cdot, t)$ is 1-periodic in each coordinate direction for all $t \in [0, T)$ and

$$\varphi_t + \sum_{j=1}^d A_j^* \varphi_{x_j} - A_0^* \varphi = 0, \quad \text{in } \mathbb{R}^d \times (0, T), \quad (3.12)$$

$$\varphi(x, T) = \Phi(x), \quad x \in \mathbb{R}^d, \quad (3.13)$$

where Φ is an arbitrary function in $C_0^\infty(\Omega_0)$.

$$\begin{aligned} (u(T) - U(T), \Phi) &= (u, \varphi)(T) - (U, \varphi)(T) \\ &= (u_0, \varphi(0)) - (U, \varphi)(T) \\ &= (u_0, \varphi(0)) - (U, \varphi)(0) + \int_0^T \frac{d}{dt} (U, \varphi) d\tau \\ &= (u_0 - P_h u_0, \varphi(0)) - \int_0^T \{(U_t, \varphi) + (U, \varphi_t)\} d\tau. \end{aligned}$$

Since, by (3.9), for $\chi : [0, T] \rightarrow \mathcal{N}_h$,

$$\begin{aligned} \int_0^T (U_t, \varphi) \, d\tau &= \int_0^T (U_t, \varphi - \chi) \, d\tau + \int_0^T (U_t, \chi) \, d\tau \\ &= \int_0^T (U_t, \varphi - \chi) \, d\tau - \int_0^T B(U, \chi) \, d\tau \\ &= \int_0^T \{ (U_t, \varphi - \chi) + B(U, \varphi - \chi) \} \, d\tau - \int_0^T B(U, \varphi) \, d\tau, \end{aligned}$$

we obtain that

$$(u(T) - U(T), \Phi) = \Theta_M + \Theta_N + \Theta_C, \quad (3.14)$$

where

$$\begin{aligned} \Theta_M &= (u_0 - P_h u_0, \varphi(0)), \\ \Theta_N &= - \int_0^T \{ (U_t, \varphi - \chi) + B(U, \varphi - \chi) \} \, d\tau, \\ \Theta_C &= - \int_0^T \{ (U, \varphi_t) - B(U, \varphi) \} \, d\tau. \end{aligned}$$

Next, we introduce some general assumptions on \mathcal{M}_h and \mathcal{N}_h which will enable us to estimate these three terms.

Let $\Omega_0 \subset\subset \Omega_1 \subset I$, $r \geq 0$, $\ell \geq 1$, and suppose that $u_0 \in L^2_{\text{per}}(\mathbb{R}^d; \mathbb{R}^m) \cap H^r(\mathcal{D}\Omega_1; \mathbb{R}^m)$, where $\mathcal{D}\Omega_1$ denotes the domain of dependence for the set Ω_1 ; see Fig. 3.1.

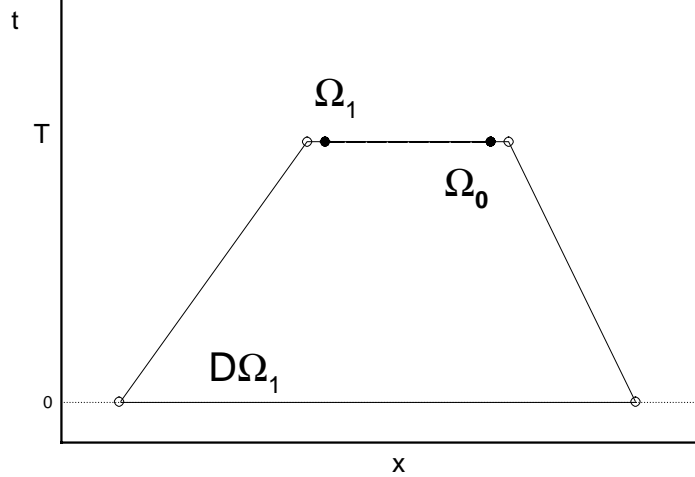


FIG. 3.1. Example of the domain of smoothness of $u(T)$, Ω_0 , and of a domain $\Omega_1 \supset\supset \Omega_0$ and its corresponding domain of dependence $\mathcal{D}\Omega_1$.

We adopt the following hypotheses.

- (i) **Approximation properties of \mathcal{M}_h and P_h .** There exist constants ρ_M, s_M , with $0 \leq \rho_M \leq \ell$ and $0 \leq s_M \leq r$, and \mathbb{A}_M such that, for each function Φ in $C_0^\infty(\Omega_0)$,

$$|(u_0 - P_h u_0, \varphi(0))| \leq \mathbb{A}_M h^{\rho_M + s_M} \|u_0\|_{r, \mathcal{D}\Omega_1} \|\Phi\|_{H^\ell},$$

where φ is the solution to the dual problem (3.12), (3.13) with the initial data Φ for the dual problem.

- (ii) **Residual.** Given that U is the solution to (3.9), (3.10), there exist constants ρ_N, s_N , with $0 \leq \rho_N \leq \ell$ and $0 \leq s_N \leq r$, and \mathbb{A}_N , such that for each function Φ in $C_0^\infty(\Omega_0)$ there exists $\chi \in C^1([0, T]; \mathcal{N}_h)$ with

$$\left| \int_0^T (U_t, \varphi - \chi)_h + B(U, \varphi - \chi) dt \right| \leq \mathbb{A}_N h^{\rho_N + s_N} \|u_0\|_{r, \mathcal{D}\Omega_1} \|\Phi\|_{H^\ell},$$

where φ is the solution to the dual problem (3.12), (3.13) with the initial data Φ for the dual problem.

- (iii) **Consistency.** Given that U is the solution to (3.9), (3.10), there exist constants $s_C \in (0, \infty]$ and $\mathbb{A}_C \in [0, \infty)$ such that

$$\left| \int_0^T (U, \varphi_t)_h - B(U, \varphi) dt \right| \leq \mathbb{A}_C h^{s_C} \|u_0\|_{r, \mathcal{D}\Omega_1} \|\Phi\|_{\ell, \Omega_0},$$

where φ is the solution to the dual problem (3.12), (3.13) with the data Φ for the dual problem.

The next result is a trivial consequence of the decomposition (3.14) and conditions (i) – (iii).

THEOREM 3.3. *Suppose that $u_0 \in L_{per}^2(\mathbb{R}^d; \mathbb{R}^m) \cap H^r(\mathcal{D}\Omega_1; \mathbb{R}^m)$, with $\Omega_0 \subset \subset \Omega_1 \subset I$, $r \geq 0$, and assume that conditions (i) – (iii) hold. Then, for $\ell \geq 1$, we have*

$$\|(u - U)(T)\|_{-\ell, \Omega_0} \leq C_4 h^s \|u_0\|_{r, \mathcal{D}\Omega_1},$$

where $s = \min\{\rho_M + s_M, \rho_N + s_N, s_C\}$ and $C_4 = \mathbb{A}_M + \mathbb{A}_N + \mathbb{A}_C$.

In Table 3.2 we display the parameters of the above result for some finite element methods; for each of the methods listed we have $\mathcal{M}_h = \mathcal{N}_h$ and have assumed that $\Omega_1 = I$ (so that $\mathcal{D}\Omega_1 = I$ also).

TABLE 3.2

The parameters of Theorem 3.3 for finite element methods using piecewise polynomials of degree k .

parameter	SG	SG with AD	DG
ρ_M	$\min\{k+1, \ell\}$	$\min\{k+1, \ell\}$	$\min\{k+1, \ell\}$
s_M	$\min\{k+1, r\}$	$\min\{k+1, r\}$	$\min\{k+1, r\}$
ρ_N	$\min\{k, \ell\}$	$\min\{k, \ell\}$	$\min\{k+1/2, \ell\}$
s_N	$\min\{k, r\}$	$\min\{k, r\}$	$\min\{k+1/2, r\}$
s_C	∞	γ	∞

3.3. The error estimates. Now we combine the results obtained in the previous subsections.

THEOREM 3.4. *Let u be the exact solution of problem (3.4), (3.5); and let U be the approximation defined by (3.9), (3.10) for which conditions (i) – (iii) are valid. Consider the convolution kernel $K_{\hat{H}}^{\nu, \ell 1}$ of Theorem 3.1. Let each of the components of $u(T)$ be in $H^\nu(\Omega_1)$ and let Ω_0 be such that $\Omega_0 + 2 \operatorname{supp}(K_{\hat{H}}^{\nu, \ell 1}) \subset \subset \Omega_1$.*

Then, for general regular triangulations and $\hat{H} = h^{s/(\nu+\ell)} \leq H_0$, we have

$$\|u(T) - K_{\hat{H}}^{\nu, \ell 1} \star U(T)\|_{0, \Omega_0} \leq C h^{\theta s},$$

where θ, s and C are as in Theorem 3.3 with $C_3 = C_4 \|u_0\|_{r, \mathcal{D}\Omega_1}$ and $\theta = \nu/(\nu + \ell)$.

Moreover, if the triangulation is translation invariant on a neighborhood of the support of the solution of the adjoint equation (3.12), (3.13) then, for $\hat{H} = h$,

$$\|u(T) - K_{\hat{H}}^{\nu, \ell 1} \star U(T)\|_{0, \Omega_0} \leq C h^s,$$

$$C_3 = C_4 \|u_0\|_{r+\ell, \mathcal{D}\Omega_1}.$$

Proof. The first inequality is a direct consequence of Corollary 3.2 and Theorem 3.3. The second inequality also follows from the above results and from the fact that if the triangulation is translation invariant in a neighborhood (of order $\hat{H} = h$) of the support of the solution of the adjoint problem, then we have

$$\|\partial_{\hat{H}}^\alpha (u - U)(T)\|_{-\ell, \Omega_0} \leq C_4 h^s \|\partial_{\hat{H}}^\alpha u_0\|_{r, \mathcal{D}\Omega_1}.$$

This completes the proof. \square

Some important particular cases for which $\Omega_1 = I$ (and consequently $\mathcal{D}\Omega_1 = I$) are collected in the table below; these are in fact the estimates we can actually prove. The case in which $\Omega_1 \neq I$ remains a challenging open problem.

TABLE 3.3

Orders of convergence with piecewise polynomials of degree k when the analytical solution u is in $C([0, T]; H_{per}^\nu(I))$.

triangulations	ν	SG	SG with AD	DG
general	0	θk	$\theta \alpha$	$\theta (k + 1/2)$
general	$2k + 2$	θk	$\theta \alpha$	$\theta (k + 1/2)$
locally invariant	0	k	α	$k + 1/2$
locally invariant	$2k + 2$	$2k$	α	$2k + 1$

4. Proofs.

4.1. The approximation result. In this subsection, for the sake of completeness, we sketch the proof of Theorem 3.1 following Bramble and Schatz [4].

Consider the following quantity:

$$\Theta_H = \|u - K_H^{\nu, \ell 1} \star U\|_{0, \Omega_0} \leq \Theta_{H,1} + \Theta_{H,2},$$

where

$$\begin{aligned} \Theta_{H,1} &= \|u - K_H^{\nu, \ell 1} \star u\|_{0, \Omega_0}, \\ \Theta_{H,2} &= \|K_H^{\nu, \ell 1} \star (u - U)\|_{0, \Omega_0}. \end{aligned}$$

To estimate $\Theta_{H,1}$, we denote the support of $K^{(\nu, \ell 1)}$ by \mathcal{I} , we label the Taylor polynomial of degree $\nu - 1$ of u around y by $T^\nu u(y, \cdot)$, and we put $\mathcal{R}^\nu u(y, \cdot) = u(\cdot) - T^\nu u(y, \cdot)$. We then easily deduce that

$$u(x) - K_H^{\nu, \ell 1} \star u(x) = R^\nu u(y, x) - \int_{\mathcal{I}} K^{\nu, \ell 1}(z) R^\nu u(y, x - Hz) dz,$$

by using the fact that the kernel $K_H^{\nu, \ell 1}$ reproduces polynomials of degree $\nu - 1$ by convolution. For $y = x$, the above expression becomes

$$u(x) - K_H^{\nu, \ell 1} \star u(x) = - \int_{\mathcal{I}} K^{\nu, \ell 1}(z) R^\nu u(x, x - Hz) dz,$$

and we obtain

$$\begin{aligned} \Theta_{H,1} &\leq \|K^{\nu, \ell 1}\|_{L^1(\mathbb{R}^d)} \sup_{z \in \mathcal{I}} \|R^\nu u(\cdot, \cdot - Hz)\|_{0, \Omega_0} \\ &\leq \frac{H^\nu}{\nu!} \|K^{\nu, \ell 1}\|_{L^1(\mathcal{I})} \|u\|_{\nu, \Omega_0 + H\mathcal{I}}. \end{aligned}$$

On applying the triangle inequality to the expression of $K^{\nu,\ell\mathbf{1}}$ given by (3.1), we get

$$\|K^{\nu,\ell\mathbf{1}}\|_{L^1(\mathcal{I})} \leq \sum_{\gamma \in \mathbb{Z}^d} |k_\gamma^{\nu,\ell\mathbf{1}}| \|\psi^{\ell\mathbf{1}-\alpha}(\cdot - \gamma)\|_{L^1(\mathbb{R}^d)} = \sum_{\gamma \in \mathbb{Z}^d} |k_\gamma^{\nu,\ell\mathbf{1}}| = C_1,$$

since $\|\psi^{\ell\mathbf{1}-\alpha}(\cdot - \gamma)\|_{L^1(\mathbb{R}^d)} = 1$. This implies that

$$\Theta_{H,1} \leq \frac{H^\nu}{\nu!} C_1 \|u\|_{\nu, \Omega_0 + H\mathcal{I}} \leq \frac{H^\nu}{\nu!} C_1 \|u\|_{\nu, \Omega_1}.$$

Now, let us sketch the procedure to estimate $\Theta_{h,2}$. Take a set $\Omega_{1/2}$ such that, for all $H \leq H_0$,

$$\Omega_0 + \text{supp}(K_h^{\nu,\ell}) \subset \Omega_{1/2},$$

$$\Omega_{1/2} + \text{supp}(K_h^{\nu,\ell}) \subset \Omega_1.$$

Then, setting $e(x) = u(x) - U(x)$, we get

$$\Theta_{H,2} = \|K_H^{\nu,\ell\mathbf{1}} \star e\|_{0,\Omega_0} \leq C_2 \sum_{|\alpha| \leq \ell} \|D^\alpha(K_H^{\nu,\ell\mathbf{1}} \star e)\|_{-\ell, \Omega_{1/2}},$$

where C_2 depends solely on Ω_0 , $\Omega_{1/2}$, d , ν , and ℓ , by Lemma 4.2 in Bramble and Schatz [4]. This is the significant step that allows us to pass from the L^2 -norm to a negative-order Sobolev norm.

Next, we exploit the fact that the kernel $K_H^{\nu,\ell\mathbf{1}}$ is a linear combination of B-splines given by (3.1); this is the only place in this proof where properties of B-splines are used. Thus, by the property (3.2) we have that

$$D^\alpha(K_H^{\nu,\ell\mathbf{1}} \star e) = K_H^{\nu,\ell\mathbf{1};\alpha} \star \partial_H^\alpha e,$$

where

$$K_H^{\nu,\ell\mathbf{1};\alpha}(y) = \sum_{\gamma \in \mathbb{Z}^d} k_\gamma^{\nu,\ell\mathbf{1}} \psi^{(\ell\mathbf{1}-\alpha)}(y - \gamma).$$

This implies that

$$\Theta_{H,2} \leq C_2 \sum_{|\alpha| \leq \ell} \|K_H^{\nu,\ell\mathbf{1};\alpha} \star \partial_H^\alpha e\|_{-\ell, \Omega_{1/2}} \leq C_2 \sum_{|\alpha| \leq \ell} \|K_H^{\nu,\ell\mathbf{1};\alpha}\|_{L^1(\mathbb{R}^d)} \|\partial_H^\alpha e\|_{-\ell, \Omega_1}.$$

Finally, since $\|K_H^{\nu,\ell\mathbf{1};\alpha}\|_{L^1(\mathbb{R}^d)} = \|K^{\nu,\ell\mathbf{1};\alpha}\|_{L^1(\mathbb{R}^d)} \leq C_1$, we get

$$\Theta_{H,2} \leq C_1 C_2 \sum_{|\alpha| \leq \ell} \|\partial_H^\alpha e\|_{-\ell, \Omega_1}.$$

This completes the proof of Theorem 3.1.

4.2. The conditions (i)–(iii) for some finite element methods. In this subsection, we justify the results displayed in Table 3.2.

a. The SG method. Let us begin by considering property (i). For the L^2 -projection, it is well known that $\rho_M = \min\{k+1, \ell\}$ and that $s_M = \min\{k+1, r\}$ for regular triangulations. Next, let us consider property (iii). A simple calculation shows that we can take $\mathbb{A}^C = 0$; this allows us to take $s_C = \infty$.

For property (ii), we proceed as follows:

$$\begin{aligned} \Theta &= \int_0^T \{(U_t, \varphi - \chi)_h + B(U, \varphi - \chi)\} dt \\ &= \int_0^T \{((U - u)_t + \mathcal{A}(U - u), \varphi - \chi)_h\} dt. \end{aligned}$$

Taking χ as the L^2 -projection of φ , we get (ii) with $\rho_N = \min\{k, \ell\}$, $s_N = \min\{k, r\}$ and $\Omega_1 = I$.

b. The SG with artificial diffusion. For this method, we only have to focus on property (iii). We have,

$$\begin{aligned}\Theta &= \int_0^T \{(U, \varphi_t)_h - B(U, \varphi)\} dt = \int_0^T h^\gamma (\nabla U, \nabla \varphi) dt \\ &\leq h^\gamma \|\nabla U\|_{L^2(0,T;L^2(I))} \|\nabla \varphi\|_{L^2(0,T;L^2(I))}.\end{aligned}$$

This means that property (iii) is satisfied with $s_C = \gamma$.

c. The DG method. For properties (i) and (iii), we proceed as in the two methods discussed above. The verification of property (ii) requires a more delicate argument. For a function W whose components are in $H_{per,h}^1(\mathbb{R}^d; \mathbb{R}^m)$, we set $\llbracket \mathcal{A} W \rrbracket(x) = \mathcal{A} \cdot n^+ W^+(x) + \mathcal{A} \cdot n^- W^-(x)$ for every $x \in e$, where $W^\pm(x) = \lim_{z \downarrow 0} W(x - z n^\pm)$ and n^\pm is orthogonal to the face e of the element K at x . With this notation, we can write that

$$\begin{aligned}\Theta &= \int_0^T \{(U_t, \varphi - \chi)_h + B(U, \varphi - \chi)\} dt \\ &= \int_0^T \left\{ (U_t + \mathcal{A}U, \varphi - \chi)_h + \sum_{e \in \mathcal{E}} \langle \llbracket \mathcal{A}U \rrbracket, \varphi - \chi \rangle_e \right\} dt,\end{aligned}$$

where we obtained the last step after a simple integration by parts; by \mathcal{E} we denote the collection of faces e of the elements K of the triangulation $\mathcal{T}_{h,I}$.

Now, taking χ to be the L^2 -projection of φ onto $\mathcal{M}_h = \mathcal{N}_h$, we get,

$$\begin{aligned}\Theta &= \int_0^T \left\{ \sum_{e \in \mathcal{E}} \langle \llbracket \mathcal{A}U \rrbracket, \varphi - \chi \rangle_e \right\} dt \\ &\leq \left\{ \int_0^T \sum_{e \in \mathcal{E}} \|\llbracket \mathcal{A}U \rrbracket\|_{e,0}^2 dt \right\}^{1/2} \left\{ \int_0^T \sum_{e \in \mathcal{E}} \|\varphi - \chi\|_{e,0}^2 dt \right\}^{1/2}.\end{aligned}$$

This implies that property (ii) is satisfied with $s_N = \min\{k + 1/2, \ell\}$ for regular triangulations. It remains to obtain an estimate of the first term on the right; following Johnson and Pitkäranta [14] or Cockburn [6], it is easy to prove that (iii) is satisfied with $\rho_N = \min\{k + 1/2, r\}$ and $\Omega_1 = I$.

5. Numerical experiments. In this section, we validate our theoretical results with an emphasis on the case in which the doubling of the order of convergence is achieved. We also explore the performance of the post-processing technique in situations not predicted by our analysis; thus, we display the L^∞ -errors in all our experiments, including an example of a linear convection-diffusion equation and an example of a non-linear convection equation.

We consider the discontinuous Galerkin method with polynomials of degree k (denoted by P^k) and use a third order Runge-Kutta method to discretize in time; the time step Δt is chosen small enough so that spatial errors dominate. Results for P^1 to P^4 are shown.

The L^∞ error measures the maximum numerical error at the six Gaussian points in each element for all elements. The L^2 error is computed by the six-point Gaussian rule in each element.

Example 5.1. A linear scalar convection equation with smooth solution on the domain $I = [0, 2\pi]$:

$$u_t + u_x = 0, \quad \text{in } I \times (0, T); \quad u(x, 0) = \sin(x) \quad x \in I, \quad (5.1)$$

with periodic boundary conditions. The errors are computed at $T = 12.5$ which is about 2 periods in time.

In Table 5.1, we show the numerical errors for this problem. We can clearly see that both the L^2 and L^∞ error for P^k elements is of $(k+1)$ -th order before post-processing and of at least $(2k+1)$ -th order after post-processing. This is consistent with our theoretical results.

TABLE 5.1
Example 5.1, $u_t + u_x = 0$, smooth solution.

mesh	Before postprocessing				After postprocessing			
	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	3.29E-02	—	5.81E-02	—	3.01E-02	—	4.22E-02	—
20	5.63E-03	2.55	1.06E-02	2.45	3.84E-03	2.97	5.44E-03	2.96
40	1.16E-03	2.28	2.89E-03	1.88	4.79E-04	3.00	6.78E-04	3.01
80	2.72E-04	2.09	8.08E-04	1.84	5.97E-05	3.00	8.45E-05	3.00
160	6.68E-05	2.03	2.13E-04	1.93	7.45E-06	3.00	1.05E-05	3.00
320	1.66E-05	2.01	5.45E-05	1.96	9.30E-07	3.00	1.32E-06	3.00
P^2								
10	8.63E-04	—	2.86E-03	—	2.52E-04	—	3.57E-04	—
20	1.07E-04	3.01	3.69E-04	2.95	5.96E-06	5.40	8.41E-06	5.41
40	1.34E-05	3.00	4.63E-05	3.00	1.53E-07	5.29	2.16E-07	5.28
80	1.67E-06	3.00	5.78E-06	3.00	4.22E-09	5.18	5.97E-09	5.18
160	2.09E-07	3.00	7.23E-07	3.00	1.27E-10	5.06	1.80E-10	5.06
P^3								
10	3.30E-05	—	9.59E-05	—	1.64E-05	—	2.31E-05	—
20	2.06E-06	4.00	6.07E-06	3.98	7.07E-08	7.85	1.00E-07	7.85
40	1.29E-07	4.00	3.80E-07	4.00	2.91E-10	7.92	4.15E-10	7.91
50	5.29E-08	4.00	1.56E-07	4.00	5.03E-11	7.87	7.24E-11	7.83
P^4								
10	1.02E-06	—	2.30E-06	—	1.98E-06	—	2.81E-06	—
20	3.21E-08	5.00	7.30E-08	4.98	2.20E-09	9.82	3.11E-09	9.82
30	4.23E-09	5.00	9.66E-09	4.99	4.34E-11	9.68	6.66E-11	9.48

In Fig. 5.1, we plot the errors of the numerical solution before and after post-processing for P^2 and P^3 with 20 elements. We can clearly see that the errors before post-processing are highly oscillatory, and the post-processing gets rid of the oscillation in the error and greatly reduces its magnitude.

In Fig. 5.2, we plot the errors, in absolute value and in logarithmic scale, of the numerical solution before and after post-processing for P^2 , with 10, 20, 40, 80, and 160 elements. We can clearly see that the post-processed errors are less oscillatory and much smaller in magnitude, and approximately third and fifth order accuracy for the pre-processed and post-processed errors, respectively, measured by the spacing

between the errors when the number of elements doubles.

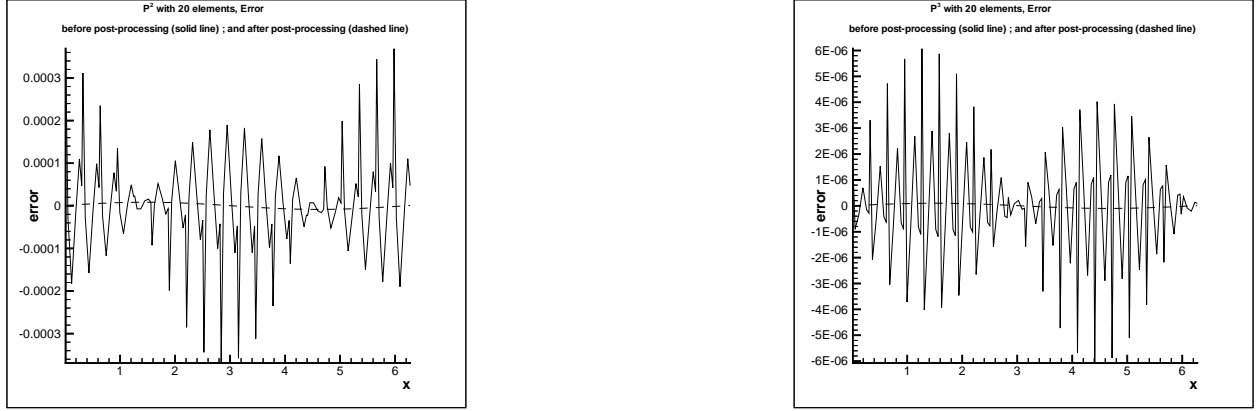


FIG. 5.1. The errors before and after post-processing for 20 elements: P^2 (left) and P^3 (right).

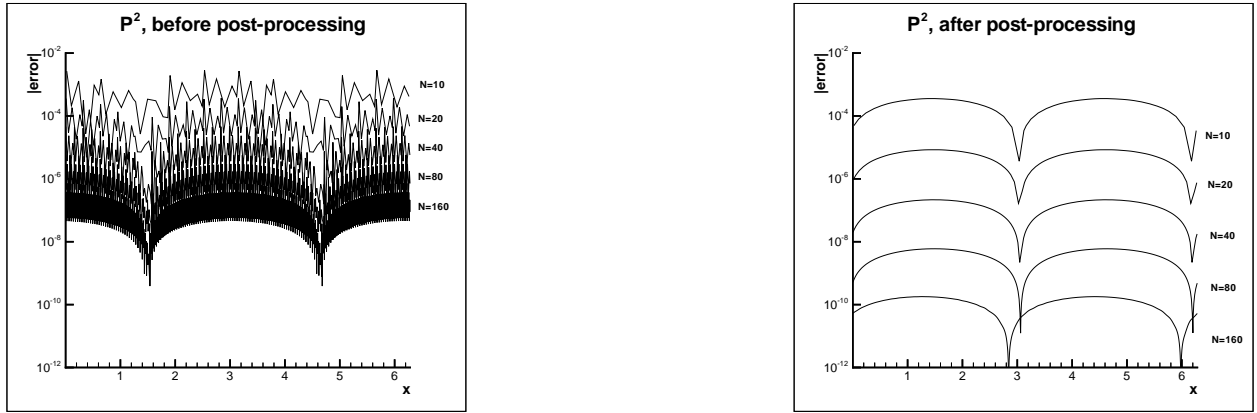


FIG. 5.2. The errors in absolute value and in logarithmic scale, for P^2 with $N=10, 20, 40, 80$, and 160 elements. Before post-processing (left) and after post-processing (right).

Example 5.2. A linear scalar convection diffusion equation with smooth solution on the domain $I = [0, 2\pi]$:

$$u_t + u_x = u_{xx}, \quad \text{in } I \times (0, T) \quad u(x, 0) = \sin(x), \quad x \in I, \quad (5.2)$$

with periodic boundary conditions. The errors are computed at $T = 2$, using the local discontinuous Galerkin method [8], with alternating left and right fluxes for u and the auxiliary variable q which approximates u_x (formula (2.9) in [8]).

Although not proven in this paper, we expect the same accuracy result to hold for the post-processed solution as in Example 5.1. In Table 5.2, we show the results for this problem. We can clearly see that the L^∞ errors for P^k elements are of $(k + 1)$ -th order before post-processing and of at least $(2k + 1)$ -th order after post-processing, both for the solution u and for the auxiliary variable q which approximates u_x . The results for the L^2 errors are similar and are not shown to save space.

Example 5.3. The same linear scalar convection equation (5.1) with the same initial condition, except that now $I = [0, 5]$. The solution now has a *discontinuity* at $x = 0$ (or $x = 5$) and this discontinuity moves in time with the characteristic speed 1. We compute the errors at $T = 12.5$, that is, after 2.5 periods in time.

TABLE 5.2
Example 5.2, $u_t + u_x = u_{xx}$, smooth solution.

	Before postprocessing				After postprocessing			
	u		q for u_x		u		q for u_x	
mesh	L^∞ error	order	L^∞ error	order	L^∞ error	order	L^∞ error	order
P^1								
10	6.74E-03	—	5.82E-03	—	1.19E-03	—	1.18E-03	—
20	1.82E-03	1.89	1.71E-03	1.77	1.34E-04	3.15	1.41E-04	3.07
40	4.68E-04	1.96	4.56E-04	1.91	1.56E-05	3.05	1.69E-05	3.06
80	1.19E-04	1.98	1.17E-04	1.96	1.46E-06	3.02	2.07E-06	3.03
160	3.00E-05	1.99	2.98E-05	1.98	2.32E-07	3.03	2.57E-07	3.01
320	7.52E-06	1.99	7.50E-06	1.99	2.87E-08	3.01	3.20E-08	3.01
P^2								
10	3.97E-04	—	3.38E-04	—	2.93E-05	—	2.96E-05	—
20	5.01E-05	2.99	4.61E-05	2.87	5.43E-07	5.75	5.46E-07	5.76
40	6.25E-06	3.00	6.02E-06	2.94	1.04E-08	5.71	1.05E-08	5.70
80	7.83E-07	3.00	7.68E-07	2.97	2.19E-10	5.57	2.26E-10	5.54
160	9.78E-08	3.00	9.69E-08	2.99	5.31E-12	5.37	5.63E-12	5.33
P^3								
10	1.30E-05	—	1.13E-05	—	3.09E-06	—	3.09E-06	—
20	8.23E-07	3.98	7.73E-07	3.86	1.32E-08	7.87	1.32E-08	7.87
40	5.14E-08	4.00	4.99E-08	3.95	5.31E-11	7.96	5.31E-11	7.96
P^4								
10	3.11E-07	—	2.93E-07	—	3.79E-07	—	3.80E-07	—
20	9.89E-09	4.97	9.54E-09	4.94	4.19E-10	9.82	4.19E-10	9.82
30	1.30E-09	5.00	1.27E-09	4.98	7.11E-12	10.05	7.11E-12	10.05

The discontinuity at this time is located at $x = 2.5$. The errors shown in Table 5.3 are calculated within the “smooth region” $[0, 1] \cup [4, 5]$, at distance 1.5 away from the discontinuity, namely *excluding* the interval $1 < x < 4$.

The theory in this paper would only guarantee $(k + 1)$ -th order accuracy for P^k elements after post-processing since our estimates hold for $\mathcal{D}\Omega_1 = I$ only and the initial condition displays a discontinuity. However, Table 5.3 shows that both the L^2 errors and the L^∞ errors are still at least $(2k + 1)$ -th order accurate for P^k elements after post-processing. This indicates that it is reasonable to expect that a similar result with a domain $\mathcal{D}\Omega_1$ *excluding* the discontinuity should hold, see Fig. 5.4.

In Fig. 5.3 we plot the errors, in absolute value and in logarithmic scale, of the numerical solution before and after post-processing for P^2 , with 10, 20, 40, 80 and 160 elements. We can clearly see that the post-processed errors are less oscillatory and much smaller in magnitude away from the discontinuity.

TABLE 5.3
Example 5.3, $u_t + u_x = 0$, discontinuous solution, errors in smooth regions

	Before postprocessing				After postprocessing			
mesh	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	2.02E-02	—	6.46E-02	—	1.76E-02	—	2.80E-02	—
20	4.37E-03	2.21	1.21E-02	2.41	3.96E-03	2.15	1.18E-02	1.24
40	6.63E-04	2.72	1.89E-03	2.69	2.69E-04	3.88	6.77E-04	4.12
80	1.58E-04	2.07	5.24E-04	1.85	2.78E-05	3.27	4.31E-05	3.97
160	3.92E-05	2.01	1.36E-04	1.94	3.49E-06	3.00	5.31E-06	3.02
320	9.80E-06	2.00	3.47E-05	1.97	4.37E-07	3.00	6.63E-07	3.00
P^2								
10	4.53E-03	—	1.08E-02	—	3.74E-03	—	1.15E-02	—
20	4.96E-04	3.19	1.98E-03	2.45	3.02E-04	3.63	1.07E-03	3.42
40	8.80E-06	5.82	2.51E-05	6.30	4.03E-06	6.23	2.74E-05	5.29
80	8.97E-07	3.29	2.91E-06	3.11	1.74E-09	11.18	1.32E-08	11.02
160	1.12E-07	3.00	3.64E-07	3.00	5.09E-11	5.09	8.75E-11	7.23
P^3								
10	2.87E-03	—	1.24E-02	—	7.76E-04	—	1.81E-03	—
20	1.97E-04	3.87	1.03E-03	3.60	6.91E-06	6.81	2.92E-05	5.95
40	1.36E-06	7.18	7.21E-06	7.15	3.51E-08	7.62	1.88E-07	7.27
80	3.03E-09	8.81	1.01E-08	9.47	2.18E-11	10.65	6.89E-11	11.42
P^4								
10	1.93E-03	—	6.32E-03	—	1.36E-03	—	2.91E-03	—
20	9.79E-05	4.30	5.42E-04	3.54	1.15E-07	13.53	8.37E-07	11.76
40	5.86E-07	7.39	4.70E-06	6.85	3.46E-11	11.70	2.11E-10	11.96

Example 5.4. A scalar nonlinear Burgers' equation with continuous and discontinuous solutions on the domain $I = [0, 2\pi)$:

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0, \quad \text{in } I \times (0, T); \quad u(x, 0) = \frac{1}{2} + \sin(x), \quad x \in I; \quad (5.3)$$

with periodic boundary conditions. The errors at $T = 0.5$, when the solution is still smooth, are given in Table 5.4. It seems that in general, post-processed errors are still smaller, although the asymptotic orders seem to show up later than for the linear case, as the mesh is refined. We remark that the theory in this paper does not cover this nonlinear problem.

In Fig. 5.5, we plot the errors of the numerical solution before and after post-processing for P^2 and P^3 with 20 elements. From Table 5.4 we can see that in both situations the errors after post-processing are actually larger than the errors before post-processing. Note in Fig. 5.5 that near the middle region, the oscillations in the errors are not “uniform”, apparently due to nonlinear effects, hence the post-processing actually gives larger errors. Fortunately, for a larger number of elements the post-processing begins to be effective and the errors after post-processing do become smaller, see Table 5.4 and the following Fig. 5.6.

In Fig. 5.6 we plot the errors, in absolute value and in logarithmic scale, of the numerical solution before and after post-processing for P^2 , with 10, 20, 40, 80, 160 and 320 elements. We can clearly see that the post-processed errors are less oscillatory and much smaller in magnitude, especially for a large number of elements. However, notice that due to non-linear effects not all oscillations in the errors have been removed by the post-processing, especially for a large number of elements.

TABLE 5.4
Example 5.4, Burgers' equation with smooth solution.

	Before postprocessing				After postprocessing			
mesh	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	1.95E-02	—	8.87E-02	—	1.37E-02	—	3.99E-02	—
20	5.31E-03	1.88	2.77E-02	1.68	1.63E-03	3.07	6.47E-03	2.63
40	1.33E-03	2.00	7.55E-03	1.87	1.28E-04	3.67	5.55E-04	3.54
80	3.33E-04	1.99	1.95E-03	1.95	1.03E-05	3.63	4.17E-05	3.73
160	8.37E-05	1.99	4.99E-04	1.97	1.12E-06	3.20	4.21E-06	3.31
320	2.10E-05	2.00	1.26E-04	1.98	1.42E-07	2.98	5.69E-07	2.89
P^2								
10	3.46E-03	—	1.93E-02	—	1.12E-02	—	3.37E-02	—
20	4.81E-04	2.85	3.57E-03	2.43	9.25E-04	3.59	3.47E-03	3.28
40	8.00E-05	2.59	6.22E-04	2.52	3.63E-05	4.67	1.58E-04	4.46
80	1.30E-05	2.62	1.20E-04	2.37	8.43E-07	5.43	3.93E-06	5.33
160	2.04E-06	2.67	1.98E-05	2.61	1.67E-08	5.66	8.51E-08	5.53
320	3.06E-07	2.73	3.02E-06	2.71	3.60E-10	5.53	1.85E-09	5.52
P^3								
10	4.33E-04	—	2.24E-03	—	1.12E-02	—	3.35E-02	—
20	4.16E-05	3.38	2.00E-04	3.48	8.08E-04	3.80	3.03E-03	3.46
40	2.43E-06	4.10	1.74E-05	3.53	2.06E-05	5.30	9.42E-05	5.01
80	1.46E-07	4.06	1.04E-06	4.07	1.96E-07	6.72	1.01E-06	6.54
160	1.03E-08	3.82	6.72E-08	3.95	1.10E-09	7.47	5.94E-09	7.41
P^4								
10	1.75E-04	—	8.25E-04	—	1.15E-02	—	3.36E-02	—
20	4.19E-06	5.39	2.45E-05	5.07	7.63E-04	3.91	2.82E-03	3.58
40	1.70E-07	4.62	1.04E-06	4.55	1.48E-05	5.69	6.82E-05	5.37
50	6.45E-08	4.36	4.40E-07	3.87	3.09E-06	7.03	1.52E-05	6.74

Next, we compute the solution at $T = 2$, that is, *after* the shock has developed. We measure the errors on the smooth region 0.5π away from the discontinuity and show the results in Table 5.5. The codes ran stably only for P^1 and P^2 ; hence only these two cases are shown.

In order to stabilize the algorithm in the presence of shocks, we apply a TVB (total variation bounded) limiter with $M = 3$, see [7]. This limiter has no effect on the numerical solution at $T = 0.5$ when the solution is still smooth, but allows the scheme to run stably for P^3 and P^4 after the shock develops. We again

TABLE 5.5
Example 5.4, Burgers' equation with discontinuous solution.

	Before postprocessing				After postprocessing			
mesh	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	8.70E-03	—	3.56E-02	—	6.79E-03	—	1.99E-02	—
20	3.05E-04	4.83	1.47E-03	4.60	2.23E-04	4.93	8.61E-04	4.53
40	1.70E-05	4.16	8.14E-05	4.18	1.09E-05	4.36	2.25E-05	5.26
80	3.71E-06	2.20	2.07E-05	1.97	1.37E-06	2.99	2.86E-06	2.97
160	8.65E-07	2.10	4.67E-06	2.15	1.63E-07	3.07	3.43E-07	3.06
320	2.17E-07	2.00	1.19E-06	1.97	2.05E-08	3.00	4.31E-08	2.99
P^2								
10	6.26E-03	—	3.29E-02	—	1.57E-03	—	7.05E-03	—
20	2.77E-04	4.50	1.44E-03	4.52	5.47E-05	4.84	1.52E-04	5.54
40	2.03E-05	3.77	1.68E-04	3.10	6.88E-06	2.99	2.62E-05	2.53
80	2.30E-06	3.14	2.17E-05	2.95	8.39E-07	3.03	4.61E-06	2.51
160	4.23E-07	2.44	4.75E-06	2.19	1.16E-07	2.86	7.95E-07	2.54
320	6.12E-08	2.79	7.77E-07	2.61	1.41E-08	3.04	1.29E-07	2.62

measure the errors on the smooth region 0.5π away from the discontinuity and show the result in Table 5.6.

In Fig. 5.7 we plot the errors, in absolute value and in logarithmic scale, of the numerical solution before and after post-processing for P^2 with a TVB limiter, at $t = 2$, with 10, 20, 40, 80, 160 and 320 elements. We can clearly see that the post-processed errors are less oscillatory and much smaller in magnitude, especially for a large number of elements, away from the discontinuity. Again, notice that not all oscillations in the errors have been removed by the post-processing, especially for a large number of elements, due to non-linear effects.

Example 5.5. A linear system with a smooth solution in the domain $I = [0, 2\pi)$:

$$\begin{cases} u_t + v_x = 0 \\ v_t + u_x = 0 \end{cases} \quad \text{in } I \times (0, T), \quad \begin{cases} u(x, 0) = \sin(x) \\ v(x, 0) = 0 \end{cases} \quad x \in I, \quad (5.4)$$

with periodic boundary conditions. The errors are computed at $T = 12.5$ which is about 2 periods in time.

In Table 5.7, we show the results for this problem. The errors are the combined ones of u and v . We can clearly see that both L^2 and L^∞ errors for P^k elements are $(k + 1)$ -th order before post-processing and at least $(2k + 1)$ -th order after post-processing. In fact, the errors are very similar to the scalar case in Example 5.1. This is consistent with our theoretical results.

Notice that this example and the next one with discontinuous solution for linear systems indicate that the method is very suitable for long time simulation of linear systems as the post-processing needs to be performed only at the final time. Examples include aeroacoustic problems when linear Euler equations must be solved for a long time to propagate the pressure waves.

Example 5.6. The same linear system (5.4) with the same initial condition, except that now $0 \leq x < 5$ and the boundary condition is 5-periodic. The solution now has a discontinuity at $x = 0$ (or $x = 5$) and this

TABLE 5.6
Example 5.4, Burgers' equation with discontinuous solution. TVB limiters.

	Before postprocessing				After postprocessing			
mesh	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	1.26E-03	—	4.44E-03	—	1.02E-03	—	2.28E-03	—
20	1.16E-04	3.44	4.38E-04	3.34	1.01E-04	3.33	1.94E-04	3.55
40	1.72E-05	2.76	8.33E-05	2.40	1.09E-05	3.22	2.26E-05	3.11
80	3.72E-06	2.20	2.08E-05	2.00	1.37E-06	2.99	2.87E-06	2.98
160	8.73E-07	2.09	4.75E-06	2.13	1.63E-07	3.07	3.44E-07	3.06
320	2.17E-07	2.01	1.19E-06	2.00	2.05E-08	3.00	4.32E-08	2.99
P^2								
10	1.03E-02	—	5.83E-02	—	1.99E-03	—	7.66E-03	—
20	4.22E-04	4.60	3.36E-03	4.12	5.32E-05	5.22	1.50E-04	5.68
40	2.23E-05	4.24	1.99E-04	4.08	6.87E-06	2.95	2.53E-05	2.53
80	2.29E-06	3.28	2.16E-05	3.20	8.39E-07	3.03	4.61E-06	2.50
160	4.21E-07	2.44	4.72E-06	2.19	1.16E-07	2.86	7.95E-07	2.54
320	6.10E-08	2.79	7.75E-07	2.61	1.41E-08	3.04	1.29E-07	2.62
P^3								
10	9.98E-04	—	6.64E-03	—	3.45E-03	—	1.35E-02	—
20	1.47E-04	2.76	1.38E-03	2.20	9.35E-06	8.52	5.18E-05	8.03
40	4.92E-07	8.22	5.43E-06	7.99	2.92E-08	8.32	2.08E-07	7.96
80	8.58E-10	9.16	1.43E-08	8.57	3.71E-10	6.30	8.73E-10	7.90
P^4								
10	3.86E-01	—	9.90E-01	—	2.28E-01	—	3.78E-01	—
20	1.08E-01	1.84	2.16E-01	2.20	5.28E-02	2.11	1.34E-01	1.50
40	1.89E-03	5.83	2.25E-02	3.26	3.88E-04	7.09	3.97E-03	5.07
80	8.08E-08	14.52	6.11E-07	15.17	1.46E-08	14.70	7.42E-08	15.71

discontinuity moves in time with the characteristic speeds ± 1 . We compute the errors at $t = 12.5$, after 2.5 periods in time. The two discontinuities at this time are both located at $x = 2.5$. The errors shown in Table 5.8 are calculated within the “smooth region” that lies a distance 1.5 away from the discontinuities, namely *excluding* the interval $1 < x < 4$.

The theory in this paper would only guarantee $(k + 1)$ -th order accuracy for P^k elements after post-processing, since when we take $\mathcal{D}\Omega_1 = I$, the initial condition has a discontinuity in this set. However, Table 5.8 shows that both the L^2 errors and the L^∞ errors are still at least $(2k + 1)$ -th order accurate for P^k elements after post-processing.

In fact, the behavior of the errors is very similar to that observed in the scalar case in Example 5.3. This is not really surprising since our linear system is equivalent to the following two decoupled scalar equations:

$$\begin{cases} (u - v)_t - (u - v)_x = 0 \\ (u + v)_t + (u + v)_x = 0 \end{cases} \quad \text{in } I \times (0, T), \quad \begin{cases} (u - v)(x, 0) = \sin(x) \\ (u + v)(x, 0) = \sin(x) \end{cases} \quad x \in I.$$

TABLE 5.7
Example 5.5, linear system with smooth solution

	Before postprocessing				After postprocessing			
mesh	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	2.33E-02	—	5.20E-02	—	2.13E-02	—	4.16E-02	—
20	3.98E-03	2.55	8.55E-03	2.60	2.72E-03	2.97	5.36E-03	2.96
40	8.20E-04	2.28	1.89E-03	2.18	3.39E-04	3.00	6.74E-04	2.99
80	1.92E-04	2.09	4.77E-04	1.98	4.23E-05	3.00	8.43E-05	3.00
160	4.72E-05	2.03	1.20E-04	2.00	5.28E-06	3.00	1.05E-05	3.00
320	1.17E-05	2.01	2.99E-05	2.00	6.59E-07	3.00	1.31E-06	3.00
P^2								
10	6.10E-04	—	1.67E-03	—	1.78E-04	—	3.53E-04	—
20	7.57E-05	3.01	2.08E-04	3.00	4.22E-06	5.40	8.42E-06	5.39
40	9.46E-06	3.00	2.60E-05	3.00	1.09E-07	5.28	2.17E-07	5.28
80	1.18E-06	3.00	3.24E-06	3.00	3.11E-09	5.13	6.20E-09	5.13
160	1.48E-07	3.00	4.06E-07	3.00	8.95E-11	5.12	1.77E-10	5.13
P^3								
10	2.33E-05	—	5.73E-05	—	1.16E-05	—	2.30E-05	—
20	1.46E-06	4.00	3.61E-06	3.99	5.00E-08	7.85	9.98E-08	7.85
40	9.13E-08	4.00	2.27E-07	3.99	2.13E-10	7.88	4.25E-10	7.88
50	3.74E-08	4.00	9.29E-08	4.01	3.94E-11	7.56	7.84E-11	7.57
P^4								
10	7.24E-07	—	1.37E-06	—	1.40E-06	—	2.79E-06	—
20	2.27E-08	5.00	4.33E-08	4.99	1.56E-09	9.82	3.11E-09	9.81
30	2.99E-09	5.00	5.72E-09	4.99	3.06E-11	9.69	6.16E-11	9.67

As a consequence, the domain of dependence $\mathcal{D}\Omega_1$ does *not* include the discontinuity of the initial condition; see Fig. 5.8 (top).

On the other hand, it is interesting to point out that this doubling of the order of accuracy does *not* take place for the problem,

$$\begin{cases} u_t - u_x = v \\ v_t + v_x = -u \end{cases} \quad \text{in } I \times (0, T), \quad \begin{cases} u(x, 0) = \sin(x) \\ v(x, 0) = 0 \end{cases} \quad x \in I, \quad (5.5)$$

with periodic boundary conditions, since now the two scalar equations associated with the diagonalization of the system are *coupled* through zero-order terms; as a consequence, the domain of dependence $\mathcal{D}\Omega_1$ *always* includes the discontinuity of the initial condition; see Fig 5.8 (bottom). This is the example treated in the early work of Majda and Osher and [18] and Majda, McDonough and Osher [17]. Due to this lack of regularity of the initial condition on $\mathcal{D}\Omega_1$, post-processing with a kernel of support of order h does not yield any significant improvement; a kernel of support of order almost one is required, as predicted by our main theorem; see also Mock and Lax [19].

TABLE 5.8
Example 5.6, linear system with discontinuous solution

mesh	Before postprocessing				After postprocessing			
	L^2 error	order	L^∞ error	order	L^2 error	order	L^∞ error	order
P^1								
10	1.49E-02	—	4.00E-02	—	1.29E-02	—	4.27E-02	—
20	3.19E-03	2.22	9.35E-03	2.10	2.79E-03	2.21	7.69E-03	2.47
40	4.76E-04	2.74	1.37E-03	2.78	1.85E-04	3.91	4.82E-04	3.99
80	1.13E-04	2.08	3.04E-04	2.17	1.99E-05	3.22	4.28E-05	3.49
160	2.78E-05	2.02	7.59E-05	2.00	2.48E-06	3.00	5.31E-06	3.01
320	6.94E-06	2.00	1.90E-05	2.00	3.09E-07	3.00	6.63E-07	3.00
P^2								
10	3.41E-03	—	6.93E-03	—	2.65E-03	—	9.29E-03	—
20	3.58E-04	3.25	1.30E-03	2.42	2.22E-04	3.58	6.40E-04	3.86
40	6.30E-06	5.83	2.42E-05	5.74	2.85E-06	6.28	1.39E-05	5.52
80	6.33E-07	3.32	1.64E-06	3.88	1.23E-09	11.17	7.52E-09	10.86
160	7.91E-08	3.00	2.05E-07	3.00	3.34E-11	5.20	5.54E-11	7.08
P^3								
10	2.03E-03	—	6.43E-03	—	5.35E-04	—	1.77E-03	—
20	1.40E-04	3.86	5.41E-04	3.57	4.92E-06	6.76	2.27E-05	6.28
40	9.66E-07	7.18	3.64E-06	7.21	2.50E-08	7.62	9.60E-08	7.89
80	2.14E-09	8.82	6.00E-09	9.25	1.34E-11	10.87	4.89E-11	10.94
P^4								
10	1.38E-03	—	3.26E-03	—	9.61E-04	—	2.90E-03	—
20	6.92E-05	4.32	2.72E-04	3.58	8.09E-08	13.54	6.31E-07	12.17
40	4.14E-07	7.39	2.36E-06	6.85	2.42E-11	11.71	1.03E-10	12.58

6. Extensions and concluding remarks. In this paper, we have shown how to enhance the approximation given by a finite element method for linear hyperbolic equations by applying a simple post-processing at the very end of the computations. Our theoretical results can be easily extended to the case in which the matrices $A_j, j = 0, \dots, d$, are very smooth functions of (x, t) . To do that, it is enough to mimic the induction argument presented by Bramble and Schatz in [4].

The role of negative-order error estimates is crucial since it is the analytical tool that captures the oscillatory nature of the error. For these negative-order norms of the error, upper bounds were obtained which depend on a *global* norm of the initial data. Our numerical results suggest, however, that they should depend only on a *local* norm of the initial data. In fact, a result of this type was obtained in 1998 for finite difference schemes by Engquist and Sjögreen [10]. To obtain such a result for, say, the discontinuous Galerkin method is an challenging open problem.

Finally, let us end by pointing out that our numerical results seem to indicate that the post-processing has a positive impact on the quality of the approximate solution even if the problem is non-linear. A theoretical analysis of this case is yet another important open problem.

Acknowledgements. The authors are grateful to Paul Castillo for providing the figures in the Introduction.

REFERENCES

- [1] S. ADJERID, M. AIFFA, AND J. E. FLAHERTY, *Computational methods for singularly perturbed systems*, Singular Perturbation Concepts of Differential Equations (J. Cronin and R.E. O'Malley, eds.), AMS Proceedings of Symposia in Applied Mathematics, AMS, 1998.
- [2] S. ADJERID, M. AIFFA, AND J.E. FLAHERTY, *High-order finite element methods for singularly-perturbed elliptic and parabolic problems*, SIAM J. Appl. Math. **55** (1995), pp. 520–543.
- [3] M.Y.T. APELKRAANS, *Some properties of difference schemes for hyperbolic equations with discontinuities and a new method with almost quadratic convergence*, Tech. Report 15A, Uppsala University, Dept. of Computer Science, 1969.
- [4] J.H. BRAMBLE AND A.H. SCHATZ, *Higher order local accuracy by averaging in the finite element method*, Math. Comp. **31** (1977), 94–111.
- [5] P. Brenner, V. Thomée, and L. Wahlbin, *Besov spaces and applications to difference methods for initial value problems*, Lecture Notes in Mathematics, Vol. 434. Springer-Verlag, Berlin-New York, 1975.
- [6] B. COCKBURN, *Discontinuous Galerkin methods for convection-dominated problems*, High-Order Methods for Computational Physics (T. Barth and H. Deconink, eds.), Lecture Notes in Computational Science and Engineering, vol. 9, Springer Verlag, 1999, pp. 69–224.
- [7] B. COCKBURN AND C.-W. SHU, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II: General framework*, Math. Comp. **52** (1989), pp. 411–435.
- [8] ———, *The local discontinuous Galerkin finite element method for convection-diffusion systems*, SIAM J. Numer. Anal. **35** (1998), pp. 2440–2463.
- [9] J. Douglas, Jr., *Superconvergence in the pressure in the simulation of miscible displacement*, SIAM J. Numer. Anal. **22** (1985), pp. 962–969.
- [10] B. ENGQUIST AND B. SJÖGREEN, *The convergence rate of finite difference schemes in the presence of shocks*, SIAM J. Numer. Anal. **35** (1998), pp. 2464–2485.
- [11] R.P. FEDORENKO, *The application of high-accuracy difference schemes to the numerical solution of hyperbolic equations*, Zh. Vychisl. Mat. i Mat. Fiz. **2** (1962), pp. 1122–1128, (in Russian).
- [12] D. GOTTLIEB AND E. TADMOR, *Recovering pointwise values of discontinuous data within spectral accuracy*, Proceedings of U.S.-Israel Workshop, Progress in Scientific Computing, vol. 6, Birkhäuser Boston Inc., 1985, pp. 357–375.
- [13] B. GUSTAFSSON, H.-O. KREISS, AND J. OLIGER, *Time dependent problems and difference methods*, John Wiley & Sons, New York, 1995.
- [14] C. JOHNSON AND J. PITKÄRANTA, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp. **46** (1986), pp. 1–26.
- [15] B.S. JOVANOVIĆ, L.D. IVANOVIĆ, AND E. SÜLI, *Convergence of a finite-difference scheme for second-order hyperbolic equations with variable coefficients*, IMA J. Numer. Anal. **7** (1987), pp. 39–45.
- [16] R.B. LOWRIE, *Compact higher-order numerical methods for hyperbolic conservation laws*, Ph.D. thesis, University of Michigan, 1996.
- [17] A. MAJDA, J. McDONOUGH, AND S. OSHER, *The Fourier method for nonsmooth initial data*, Math.

- Comp. **32** (1978), pp. 1041–1081.
- [18] A. MAJDA AND S. OSHER, *Propagation of error into regions of smoothness for accurate difference approximate solutions to hyperbolic equations*, Comm. Pure and Appl. Math **30** (1977), pp. 671–705.
 - [19] M.S. MOCK AND P.D. LAX, *The computation of discontinuous solutions of linear hyperbolic equations*, Comm. Pure and Appl. Math **31** (1978), pp. 423–430.
 - [20] V. THOMÉE, *High order local approximations to derivatives in the finite element method*, Math. Comp. **31** (1977), pp. 652–660.
 - [21] L.B. WAHLBIN, *Superconvergence in Galerkin finite element methods*, Lecture Notes in Mathematics, vol. 1605, Springer Verlag, 1995.

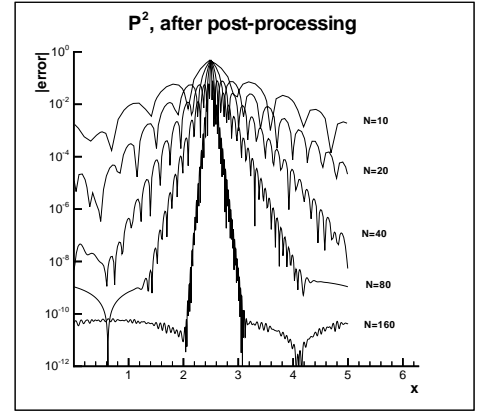
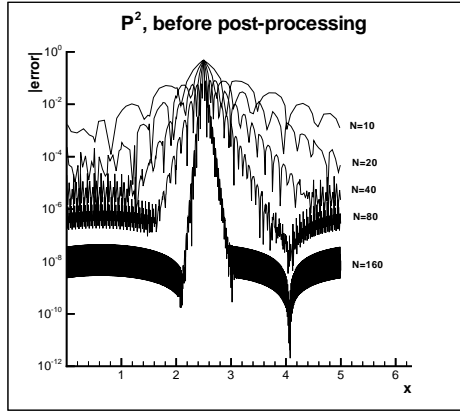


FIG. 5.3. The errors in absolute value and in logarithmic scale, for P^2 with $N=10, 20, 40, 80$ and 160 elements. Before post-processing (left) and after post-processing (right).

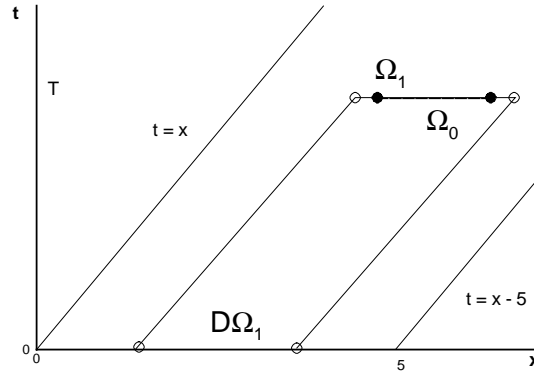


FIG. 5.4. Sketch of the domain of smoothness Ω_0 of the exact solution, the domain Ω_1 and its corresponding domain of dependence $\mathcal{D}\Omega_1$. The pattern between the discontinuities $t = x$ and $t = x - 5$ should be repeated periodically.

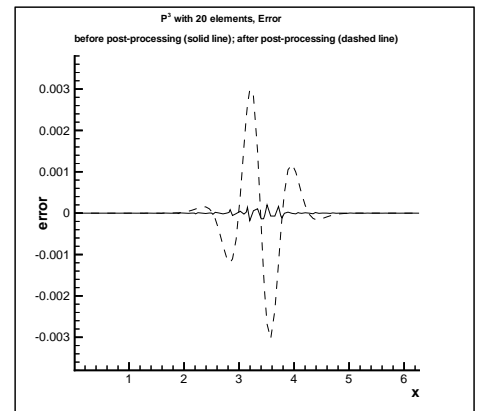
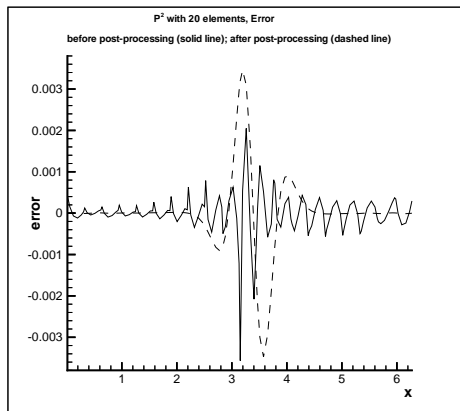


FIG. 5.5. The errors before and after post-processing for the smooth solution of Burgers equation and 20 elements: P^2 (left) and P^3 (right).

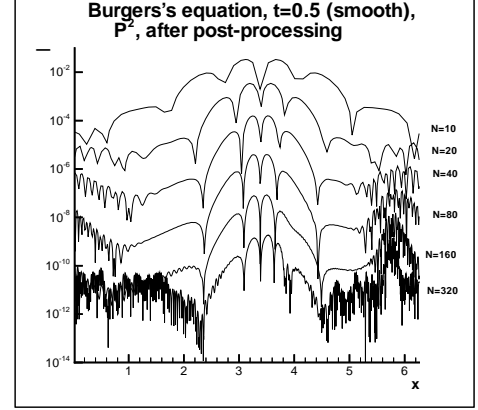
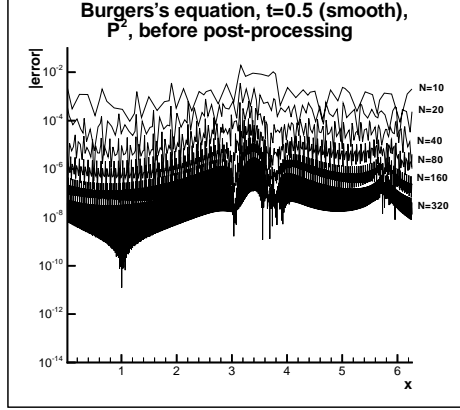


FIG. 5.6. The errors in absolute value and in logarithmic scale, for P^2 with $N=20, 40, 80, 160$ and 320 elements. Smooth solution of Burgers equation. Before post-processing (left) and after post-processing (right).

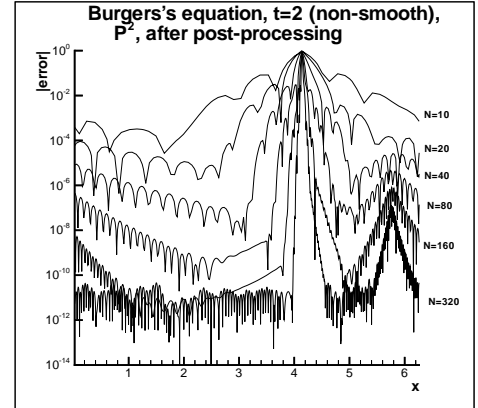
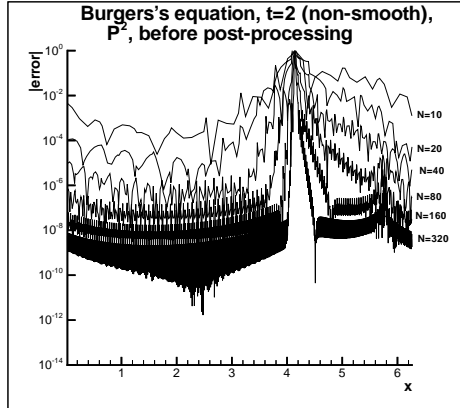


FIG. 5.7. The errors in absolute value and in logarithmic scale, for P^2 with $N=10, 20, 40, 80, 160$ and 320 elements. Discontinuous solution of Burgers equation. Before post-processing (left) and after post-processing (right).

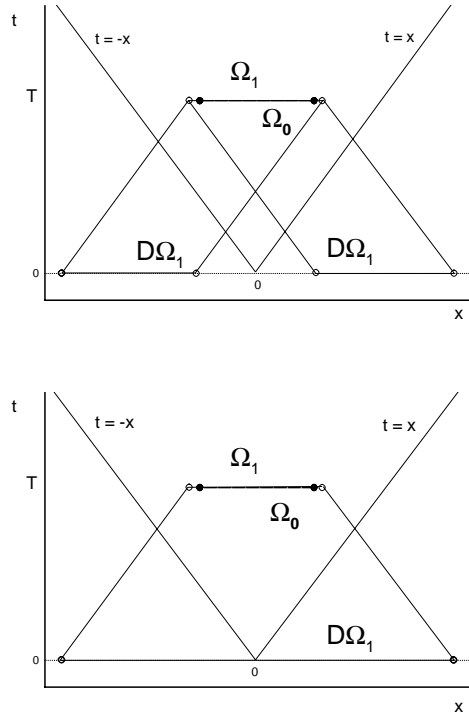


FIG. 5.8. The domain of smoothness of $u(T)$, Ω_0 , the domain $\Omega_1 \supset \supset \Omega_0$ and its corresponding domain of dependence $\mathcal{D}\Omega_1$ for the system (5.4) (top) and the system (5.5) (bottom). Note the discontinuity curves $t = |x|$.